Ethan Frenette

Mistral Text Summarization and Basic Chat AI

Due: 12/2/24

This project shows the implementation of a basic chatbot using the mistral model. This bot also has Arxiv API to provide summaries and detailed analyses of academic papers related to data science. The primary goal is to enable users to search for and explore academic content interactively. The dataset includes live query results from the API, dynamically processed to generate concise summaries and actionable insights using Mistral-7B Instruct model. Key limitations are the dependencies on GPU hardware for local model inference and the occasional repetition in responses. The chatbot has significant business value in saving researches time by automation paper exploration and analysis, making it a potential tool for academic and industrial research professionals.

I used a bundle of methods to get this to the working state that it is in, as this was a very in depth version of what was asked of in the project rubric. One of which was Data Acquisition, the api was queried to fetch academic paper data, including titles, abstracts, and metadata. Also used EDA, EDA was focused on understanding the structure of the Arxiv API’s XML response to extract relevant fields like titles and abstracts. I also had the implementation of python based chatbot to allow users to interact with the bot a little. I had also added a feature to dive deeper beyond the baseline summary the bot gives, giving a much more in depth summary of what is actually in the paper.

In conclusion, this project successfully demonstrated the integration of a language model with an API to build an interactive chatbot. Users can query the Arxiv database for papers, navigate through results, and request detailed analyses or summaries. While the chatbot effectively simplifies academic research, it is constrained by computational requirements and occasional response problems , such as repeated words after working too long. If I were to work on this longer, I would add more interactive features so that the user can have more freedom to learn about the topics, such as the chatbot giving tips on how to study the topic and such.

The chatbot offers significant business value in academic and industrial research. Researches and professionals can quickly explore academic content, saving time on manual searches and enabling faster decision-making. The chatbot could be packaged as a SaaS product, targeting universities and research institutions. The tool’s functionality could be monetized through integration with existing academic databases to offer a more personalized bot.